function of the gene in which they reside. The implication is that while normal diploid humans possess two functional copies of most genes, there are many instances in all humans where only one (or zero) functional gene copies are present. Similarly, instances where genes are overrepresented by gene duplication/amplification events are also encountered with significant frequency.

One of the key features in biological networks is functional redundancy. Normal, healthy individuals can tolerate the average load of genetic lesions because they possess, on average, two copies of every gene such that loss of one copy is inconsequential. Moreover, sets of genes often perform similar functions such that minor perturbations in specific gene functions are generally compensated for within a larger network of functional elements. While functional compensation in biological systems is a general theme, there are many instances in which specific gene loss can trigger acute disruptive events. By way of example, cancers appear to be the consequence of genetic disease in which the compound effect of multiple individual lesions is uncontrolled cell proliferation. Similarly, prescribed medicines are often specific chemical entities that are transported, metabolized and/or eliminated by very specific genes. Perturbations in these genes, while generally inconsequential under normal circumstances, can manifest as adverse events (*e.g.*, side effects) during chemical therapy.

The central aim of "personalized medicine", increasingly referred to as "precision medicine," is to merge genetic information that is specific to the patient with treatment options that are compatible with the individual's genetic profile. However, the vast potential of personalized medicine has yet to be realized. To realize this goal, there must be clinically acceptable, robust genetic diagnostic tests that can reliably determine the genetic status of relevant genes.

SUMMARY

Particular examples contemplated herein provide a method for generating a tagged DNA library comprising treating fragmented DNA with end-repair enzymes to generate fragmented end-repaired DNA; and ligating a random nucleic acid tag

sequence, and optionally a sample code sequence and/or a PCR primer sequence to the fragmented end-repaired DNA to generate the tagged DNA library.

In particular examples, the random nucleic acid tag sequence is from about 2 to about 100 nucleotides. In some examples, the present disclosure provides that the random nucleic acid tag sequence is from about 2 to about 8 nucleotides.

In certain examples, the fragmented end-repaired DNA contains blunt ends. In some examples, the blunt ends are further modified to contain a single base pair overhang.

In certain examples, the ligating comprises ligating a multifunctional adaptor module to the fragmented end-repaired DNA to generate the tagged DNA library, wherein the multifunctional adaptor molecule comprises: i) a first region comprising a random nucleic acid tag sequence; ii) a second region comprising a sample code sequence; and iii) a third region comprising a PCR primer sequence.

In additional examples, the method further comprises hybridizing a tagged DNA library with at least one multifunctional capture probe module to form a complex, wherein the multifunctional capture probe module hybridizes to a specific target region in the DNA library.

In further examples, the method further comprises isolating the tagged DNA library-multifunctional capture probe module complex.

In some examples, the method further comprises 3′-5′ exonuclease enzymatic processing of the isolated tagged DNA library-multifunctional capture probe module complex to remove the single stranded 3′ ends. In some examples, the enzyme for use in the 3′-5′ exonuclease enzymatic processing is T4 polymerase.

In particular examples, the method further comprises 5’-3’ DNA polymerase extension of the isolated tagged DNA library-multifunctional capture probe module complex from the 3’ end of the multifunctional capture probe utilizing the isolated tagged DNA library fragments as template.

In certain examples, the method further comprises joining of the multifunctional capture probe and isolated tagged DNA library fragments through the concerted action of a 5’ FLAP endonuclease, DNA polymerization and nick closure by a DNA ligase.

In further examples, the method further comprises performing PCR on the 3′-5′ exonuclease enzymatically processed complex, wherein the tail portion of the multifunctional capture probe molecule is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence.

In various examples, a method for targeted genetic analysis is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe module complex from a); c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged DNA library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; d) performing PCR on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; and e) performing targeted genetic analysis on the hybrid nucleic acid molecule from d).

In various particular examples, a method for targeted genetic analysis is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged genomic library-multifunctional capture probe module complex from a); c) performing 5′-3′ DNA polymerase extension of multifunctional capture probe utilizing the isolated tagged DNA library fragments as template; d) performing PCR on the enzymatically processed complex from c) wherein the complement of the isolated target region is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the complement of the DNA target region, the target-specific region of the multifunctional capture probe and the multifunctional capture

probe module tail sequence; and e) performing targeted genetic analysis on the hybrid nucleic acid molecule from d).

In various certain examples, a method for targeted genetic analysis is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe module complex from a); c) performing the creation of a hybrid multifunctional capture probe-isolated tagged DNA target molecule through the concerted action of a 5' FLAP endonuclease, DNA polymerization and nick closure by a DNA ligase; d) performing PCR on the enzymatically processed complex from c) wherein the of the multifunctional capture probe molecule is joined to the isolated tagged DNA target clone in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module; and e) performing targeted genetic analysis on the hybrid nucleic acid molecule from d).

In particular examples, a method for determining copy number of a specific target region is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe module complex from a); c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged DNA library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; d) performing a PCR reaction on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is replicated in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; e) performing PCR amplification of the hybrid nucleic acid in d); and e) quantitating the PCR reaction

in d), wherein the quantitation allows for a determination of copy number of the specific target region.

In certain examples, a method for determining copy number of a specific target region is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe module complex from a); c) performing 5′-3′ DNA polymerase extension of multifunctional capture probe utilizing the isolated tagged DNA library fragments as template; d) performing a PCR reaction on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is replicated in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; e) performing PCR amplification of the hybrid nucleic acid in d); and e) quantitating the PCR reaction in d), wherein the quantitation allows for a determination of copy number of the specific target region.

In further examples, a method for determining copy number of a specific target region is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe module complex from a); c) performing the creation of a hybrid multifunctional capture probe-isolated tagged DNA target molecule through the concerted action of a 5’ FLAP endonuclease, DNA polymerization and nick closure by a DNA ligase; d) performing a PCR reaction on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is replicated in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; e) performing PCR amplification

of the hybrid nucleic acid in d); and e) quantitating the PCR reaction in d), wherein the quantitation allows for a determination of copy number of the specific target region.

In additional examples, a method for targeted genetic analysis is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe hybrid module complex from a); c) performing PCR on the complex from b) to replicate a region that is 3′ relative to the sequence of the multifunctional capture probe in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of a region of the tagged DNA library sequence that is located 3′ relative to the multifunctional capture probe; and d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In particular examples, a method for targeted genetic analysis is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the genomic library; b) isolating the tagged DNA library-multifunctional capture probe hybrid module complex from a); c) performing 5′-3′ DNA polymerase extension of multifunctional capture probe utilizing the isolated tagged DNA library fragments as template, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of a region of the tagged DNA library sequence that is located 3′ relative to the multifunctional capture probe; and d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In certain examples, a method for targeted genetic analysis is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe hybrid module complex from a); c) performing the creation of a hybrid multifunctional capture probe-isolated tagged DNA target molecule through the concerted action of a 5’ FLAP endonuclease, DNA

polymerization and nick closure by a DNA ligase, wherein the hybrid nucleic acid molecule comprises the complement of the multifunctional capture probe hybrid module and a region of the tagged DNA library sequence that is located 5′ relative to the multifunctional capture probe; and d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In particular examples, a method for determining copy number of a specific target region is provided comprising: a) hybridizing a tagged DNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the DNA library; b) isolating the tagged DNA library-multifunctional capture probe hybrid module complex from a); c) performing PCR on the complex from b) to replicate a region that is 3’ relative to sequence of the multifunctional capture probe in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of a region of the tagged DNA library sequence that is located 3’ relative to the multifunctional capture probe; d) performing PCR amplification of the hybrid nucleic acid in c); and e) quantitating the PCR reaction in d), wherein the quantitation allows for a determination of copy number of the specific target region.

In various examples, the targeted genetic analysis is sequence analysis.

In particular examples, the tagged DNA library is amplified by PCR to generate an amplified tagged DNA library.

In certain examples, the DNA is from a biological sample selected from the group consisting of blood, skin, hair, hair follicles, saliva, oral mucous, vaginal mucous, sweat, tears, epithelial tissues, urine, semen, seminal fluid, seminal plasma, prostatic fluid, pre-ejaculatory fluid (Cowper's fluid), excreta, biopsy, ascites, cerebrospinal fluid, lymph, and tissue extract sample or biopsy sample.

In further examples, a tagged DNA library comprises tagged DNA sequences, wherein each tagged DNA sequence comprises: i) fragmented end-repaired DNA; ii) a random nucleotide tag sequence; iii) a sample code sequence; and iv) a PCR primer sequence.

In additional examples, a hybrid tagged DNA library comprises hybrid tagged DNA sequences for use in targeted genetic analysis, wherein each hybrid tagged DNA sequence comprises: i) fragmented end-repaired DNA; ii) a random nucleotide tag sequence; iii) a sample code sequence; iv) a PCR primer sequence; and v) a multifunctional capture probe module tail sequence.

In further examples, a multifunctional adaptor module comprises: i) a first region comprising a random nucleotide tag sequence; ii) a second region comprising a sample code sequence; and iii) a third region comprising a PCR primer sequence.

In particular examples, a multifunctional capture probe module comprises: i) a first region capable of hybridizing to a partner oligonucleotide; ii) a second region capable of hybridizing to a specific target region; and iii) a third region comprising a tail sequence. In some examples, the first region of the capture probe module is bound to a partner oligonucleotide. In some examples, the partner oligonucleotide is chemically modified.

In one example, compositions comprise a tagged DNA library, a multifunctional adaptor module and a multifunctional capture probe module.

In particular examples, a composition comprises a hybrid tagged genomic library according to the methods of the present disclosure.

In certain examples, a composition comprises a reaction mixture for performing the methods contemplated herein.

In one aspect, the present disclosure refers to a plurality of compositions for generating a captured library comprising hybrid nucleic acid molecules, the plurality of compositions comprising: (a) a first composition comprising a multifunctional adaptor module comprising: (i) a first region comprising a random nucleotide tag sequence, (ii) a second region comprising a sample code sequence, and (iii) a third region comprising a PCR primer sequence; characterized in that the multifunctional adaptor module is capable of ligating to each end of a DNA fragment, and (b) a second composition comprising a multifunctional capture probe module comprising: (i) a first region capable of hybridizing to a partner oligonucleotide, and (ii) a second region capable of hybridizing to a specific genomic target region in the DNA fragment.

In particular examples, a reaction mixture capable of generating a tagged DNA library comprises: a) fragmented DNA and b) DNA end-repair enzymes to generate fragmented end-repaired DNA.

In certain examples, a reaction mixture further comprises a multifunctional adaptor module.

In additional examples, a reaction mixture further comprises a multifunctional capture probe module.

In some examples, a reaction mixture further comprises an enzyme with 3′-5′ exonuclease activity and PCR amplification activity.

In one example, the reaction mixture comprises a FLAP endonuclease, a DNA polymerase, and DNA ligase.

In any of the foregoing examples, the DNA can be isolated genomic DNA or cDNA.

In various examples, a method for generating a tagged genomic library is provided comprising: treating fragmented genomic DNA with end-repair enzymes to generate fragmented end-repaired genomic DNA; and ligating a random nucleic acid tag sequence, and optionally a sample code sequence and/or a PCR primer sequence to the fragmented end-repaired genomic DNA to generate the tagged genomic library.

In particular examples, the random nucleic acid tag sequence is from about 2 to about 100 nucleotides.

In certain examples, the random nucleic acid tag sequence is from about 2 to about 8 nucleotides.

In additional examples, the fragmented end-repaired genomic DNA contains blunt ends.

In further examples, the blunt ends are further modified to contain a single base pair overhang.

In some examples, the ligating comprises ligating a multifunctional adaptor module to the fragmented end-repaired genomic DNA to generate the tagged genomic library, wherein the multifunctional adaptor molecule comprises: a first region comprising a random nucleic acid tag sequence; a second region comprising a sample code sequence; and a third region comprising a PCR primer sequence.

In particular examples, the methods contemplated herein comprise hybridizing a tagged genomic library with a multifunctional capture probe module to form a complex, wherein the multifunctional capture probe module hybridizes to a specific genomic target region in the genomic library.

In certain particular examples, the methods contemplated herein comprise isolating the tagged genomic library-multifunctional capture probe module complex.

In additional particular examples, the methods contemplated herein comprise 3′-5′ exonuclease enzymatic processing of the isolated tagged genomic library-multifunctional capture probe module complex to remove the single stranded 3′ ends.

In further particular examples, the enzyme for use in the 3′-5′ exonuclease enzymatic processing is T4 DNA polymerase.

In some particular examples, the methods contemplated herein comprise performing PCR on the 3′-5′ exonuclease enzymatically processed complex from the preceding claims, wherein the tail portion of the multifunctional capture probe molecule is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence.

In various examples, a method for targeted genetic analysis is provided comprising: (a) hybridizing a tagged genomic library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific genomic target region in the genomic library; (b) isolating the tagged genomic library-multifunctional capture probe module complex from a); (c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged genomic library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; (d) performing PCR on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; and (e) performing targeted genetic analysis on the hybrid nucleic acid molecule from d).

In particular examples, steps a) through d) are repeated at least about twice and the targeted genetic analysis of e) comprises a sequence alignment of the hybrid nucleic acid molecule sequences obtained from the at least two d) steps.

In further examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In some examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In various examples, a method for determining copy number of a specific genomic target region is provided comprising: (a) hybridizing a tagged genomic library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module complex selectively hybridizes to a specific genomic target region in the genomic library; (b) isolating the tagged genomic library-multifunctional capture probe module complex from a); (c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged genomic library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; (d) performing a PCR reaction on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is replicated in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; (e) performing PCR amplification of the hybrid nucleic acid molecule in d); and (f) quantitating the PCR reaction in e), wherein the quantitation allows for a determination of copy number of the specific genomic target region.

In some examples, the methods contemplated herein comprise obtaining the sequences of the hybrid nucleic acid molecules from step e).

In further examples, steps a) through e) are repeated at least about twice and a sequence alignment is performed using the hybrid nucleic acid molecule sequences obtained from the at least two e) steps.

In additional examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In certain examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In various examples, a method for determining copy number of a specific genomic target region is provided comprising: (a) hybridizing a tagged genomic library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module complex selectively hybridizes to a specific genomic target region in the genomic library; (b) isolating the tagged genomic library-multifunctional capture probe module complex from a); (c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged genomic library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; (d) performing a PCR reaction on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is replicated in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; and (e) performing PCR amplification of the hybrid nucleic acid molecule in d).

In certain examples, the methods contemplated herein comprise obtaining the sequences of the hybrid nucleic acid molecules from step e).

In particular examples, steps a) through e) are repeated at least about twice and a sequence alignment is performed using the hybrid nucleic acid molecule sequences obtained from the at least two e) steps.

In some examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In additional examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In various examples, a method for determining copy number of a specific genomic target region is provided comprising: (a) hybridizing a tagged genomic library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module comprises selectively hybridizes to a specific genomic target region in the genomic library; (b) isolating the tagged genomic library-multifunctional

capture probe module complex from a); (c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged genomic library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; (d) performing a PCR reaction on the enzymatically processed complex from c) wherein the tail portion of the multifunctional capture probe molecule is replicated in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the genomic target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; (e) performing PCR amplification of the hybrid nucleic acid molecule in d); and (f) performing targeted genetic analysis on the hybrid nucleic acid molecule from e).

In particular examples, steps a) through e) are repeated at least about twice and the targeted genetic analysis of f) comprises performing a sequence alignment of the hybrid nucleic acid molecule sequences from the at least two e) steps.

In certain examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In additional examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In various examples, a method for targeted genetic analysis is provided comprising: (a) hybridizing a tagged genomic library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific genomic target region in the genomic library; (b) isolating the tagged genomic library-multifunctional capture probe hybrid module complex from a); (c) performing 5′ to 3′ DNA polymerase extension of the multifunctional capture probe on the complex from b) to replicate a region of the captured, tagged genomic target region that is 3′ of the multifunctional capture probe in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of a region of the tagged genomic target region that is located in the 3′

direction from the location where the multifunctional capture probe hybrid module hybridizes to the genomic target region; and (d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In further examples, steps a) through c) are repeated at least about twice and the targeted genetic analysis of d) comprises a sequence alignment of the hybrid nucleic acid molecule sequences obtained from the at least two d) steps.

In some examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In particular examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In various examples, a method for determining copy number of a specific genomic target region is provided comprising: (a) hybridizing a tagged genomic library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific genomic target region in the genomic library; (b) isolating the tagged genomic library-multifunctional capture probe hybrid module complex from a); (c) performing 5' to 3' DNA polymerase extension of the multifunctional capture probe on the complex from b) to replicate a region of the captured tagged genomic target region that is 3' of the multifunctional capture probe in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of a region of the tagged genomic target region that is located in the 3'direction from the location where the multifunctional capture probe hybrid module hybridizes to the genomic target region; and (d) performing PCR amplification of the hybrid nucleic acid molecule in c); and (e) quantitating the PCR reaction in d), wherein the quantitation allows for a determination of copy number of the specific genomic target region.

In particular examples, the methods contemplated herein comprise obtaining the sequences of the hybrid nucleic acid molecules from step d).

In certain examples, steps a) through d) are repeated at least about twice and a sequence alignment of the hybrid nucleic acid molecules from the at least two d) steps.

In additional examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In further examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In some examples, the targeted genetic analysis is sequence analysis.

In particular examples, the tagged genomic library is amplified by PCR to generate an amplified tagged genomic library.

In related particular examples, the genomic DNA is from a biological sample selected from the group consisting of blood, skin, hair, hair follicles, saliva, oral mucous, vaginal mucous, sweat, tears, epithelial tissues, urine, semen, seminal fluid, seminal plasma, prostatic fluid, pre-ejaculatory fluid (Cowper's fluid), excreta, biopsy, ascites, cerebrospinal fluid, lymph, and tissue extract sample or biopsy sample.

In various examples, a tagged genomic library is provided comprising tagged genomic sequences, wherein each tagged genomic sequence comprises: fragmented end-repaired genomic DNA; a random nucleotide tag sequence; a sample code sequence; and a PCR primer sequence.

In various related examples, a tagged cDNA library is provided comprising tagged cDNA sequences, wherein each tagged cDNA sequence comprises: fragmented end-repaired cDNA; a random nucleotide tag sequence; a sample code sequence; and a PCR primer sequence.

In various particular examples, a hybrid tagged genomic library is provided comprising hybrid tagged genomic sequences for use in targeted genetic analysis, wherein each hybrid tagged genomic sequence comprises: fragmented end-repaired genomic DNA; a random nucleotide tag sequence; a sample code sequence; a PCR primer sequence; a genomic target region; and a multifunctional capture probe module tail sequence.

In various certain examples, a hybrid tagged cDNA library is provided comprising hybrid tagged cDNA sequences for use in targeted genetic analysis, wherein each hybrid tagged cDNA sequence comprises: fragmented end-repaired cDNA; a random nucleotide tag sequence; a sample code sequence; a PCR primer sequence; a cDNA target region; and a multifunctional capture probe module tail sequence.

In various certain examples, a multifunctional adaptor module is provided comprising: a first region comprising a random nucleotide tag sequence; a second region comprising a sample code sequence; and a third region comprising a PCR primer sequence.

In various additional examples, a multifunctional capture probe module is provided comprising: a first region capable of hybridizing to a partner oligonucleotide; a second region capable of hybridizing to a specific genomic target region; and a third region comprising a tail sequence.

In particular examples, the first region is bound to a partner oligonucleotide.

In particular examples, a multifunctional adaptor probe hybrid module is provided comprising: a first region capable of hybridizing to a partner oligonucleotide and capable of functioning as a PCR primer and a second region capable of hybridizing to a specific genomic target region.

In certain examples, the first region is bound to a partner oligonucleotide.

In some examples, the partner oligonucleotide is chemically modified.

In further examples, a composition comprising a tagged genomic library, a multifunctional adaptor module and a multifunctional capture probe module is provided.

In additional examples, a composition comprising a hybrid tagged genomic or c DNA library according to any of the preceding examples is provided.

In various examples, a reaction mixture for performing a method of any one of the preceding examples is provided.

In particular examples, a reaction mixture capable of generating a tagged genomic library is provided comprising: fragmented genomic DNA; and DNA end-repair enzymes to generate fragmented end-repaired genomic DNA.

In particular examples, a reaction mixture capable of generating a tagged genomic library is provided comprising: fragmented cDNA; and DNA end-repair enzymes to generate fragmented end-repaired cDNA.

In particular examples, a reaction mixture comprises a multifunctional adaptor module.

In some examples, a reaction mixture comprises a multifunctional capture probe module.

In certain examples, a reaction mixture comprises an enzyme with 3′-5′ exonuclease activity and PCR amplification activity.

In various examples, a method for DNA sequence analysis is provided comprising: obtaining one or more clones, each clone comprising a first DNA sequence and a second DNA sequence, wherein the first DNA sequence comprises targeted genomic DNA sequence and the second DNA sequence comprises a capture probe sequence; performing a paired end sequencing reaction on the one or more clones and obtaining one or more sequencing reads; and ordering or clustering the sequencing reads of the one or more clones according to the probe sequence of the sequencing reads.

In particular examples, a method for DNA sequence analysis is provided comprising: obtaining one or more clones, each clone comprising a first DNA sequence and a second DNA sequence, wherein the first DNA sequence comprises targeted genomic DNA sequence and the second DNA sequence comprises a capture probe sequence; performing a sequencing reaction on the one or more clones in which a single long sequencing read of greater than about 100 nucleotides is obtained, wherein the read is sufficient to identify both the first DNA sequence and the second DNA sequence; and ordering or clustering sequencing reads of the one or more clones according to the probe sequence of the sequencing reads.

In certain examples, the sequences of the one or more clones are compared to one or more human reference DNA sequences.

In additional examples, sequences that do not match the one or more human reference DNA sequences are identified.

In further examples, non-matching sequences are used to create a de novo assembly from the non-matching sequence data.

In some examples, the de novo assemblies are used to identify novel sequence rearrangements associated with the capture probe.

In various examples, a method for genomic copy number determination analysis is provided comprising: obtaining one or more clones, each clone comprising a first DNA sequence and a second DNA sequence, wherein the first DNA sequence comprises a random nucleotide tag sequence and a targeted genomic DNA sequence and the second DNA sequence comprises a capture probe sequence; performing a paired end sequencing reaction on the one or more clones and obtaining one or more sequencing reads; and ordering or clustering the sequencing reads of the one or more clones according to the probe sequence of the sequencing reads.

In some examples, a method for genomic copy number determination analysis is provided comprising: obtaining one or more clones, each clone comprising a first DNA sequence and a second DNA sequence, wherein the first DNA sequence comprises a random nucleotide tag sequence and a targeted genomic DNA sequence and the second DNA sequence comprises a capture probe sequence; performing a sequencing reaction on the one or more clones in which a single long sequencing read of greater than about 100 nucleotides is obtained, wherein the read is sufficient to identify both the first DNA sequence and the second DNA sequence; and ordering or clustering sequencing reads of the one or more clones according to the probe sequence of the sequencing reads.

In certain examples, the random nucleotide tag sequence is about 2 to about 50 nucleotides in length.

In further examples, the methods contemplated herein comprise analyzing all sequencing reads associated with a second read sequence by: determining the distributions of unique and redundant sequencing reads; counting the number of times a unique read is encountered; fitting a frequency distribution of the unique reads to a statistical distribution; inferring a total number of unique reads; and normalizing the total number of inferred unique read to an assumption that most human genetic loci are generally diploid.

In additional examples, an inferred copy number of one or more targeted loci are determined.

In some examples, the one or more target loci that deviate from an expected copy number value are determined.

In further examples, the one or more targeted loci of a gene are grouped together in a collection of loci and the copy number measurements from the collection of targeted loci are averaged and normalized.

In additional examples, the inferred copy number of a gene is represented by the normalized average of all the target loci representing that gene.

In certain examples, a method for generating a tagged RNA expression library is provided comprising: fragmenting a cDNA library; treating the fragmented cDNA library with end-repair enzymes to generate fragmented end-repaired cDNA; and ligating a multifunctional adapter molecule to the fragmented end-repaired c DNA to generate a tagged RNA expression library.

In particular examples, a method for generating a tagged RNA expression library is provided comprising: preparing a cDNA library from the total RNA of one or more cells; fragmenting the cDNA library; treating the fragmented cDNA with end-repair enzymes to generate fragmented end-repaired cDNA; and ligating a multifunctional adapter molecule to the fragmented end-repaired c DNA to generate a tagged RNA expression library.

In various examples, the cDNA library is an oligo-dT primed cDNA library.

In particular examples, the cDNA library is primed by random oligonucleotides comprising about 6 to about 20 random nucleotides.

In certain examples, the cDNA library is primed by random hexamers or random octamers.

In additional examples, the cDNA library is fragmented to a size of about 250 bp to about 750 bp.

In further examples, the cDNA library is fragmented to a size of about 500 bp.

In some examples, the multifunctional adaptor module comprises: a first region comprising a random nucleic acid tag sequence, and optionally; a second region

comprising a sample code sequence, and optionally a third region comprising a PCR primer sequence.

In related examples, the multifunctional adaptor module comprises a first region comprising a random nucleic acid tag sequence, a second region comprising a sample code sequence, and a third region comprising a PCR primer sequence.

In various examples, the methods contemplated herein comprise hybridizing a tagged cDNA library with a multifunctional capture probe module to form a complex, wherein the multifunctional capture probe module hybridizes to a specific target region in the cDNA library.

In some examples, the methods contemplated herein comprise isolating the tagged cDNA library-multifunctional capture probe module complex.

In particular examples, the methods contemplated herein comprise 3′-5′ exonuclease enzymatic processing of the isolated tagged cDNA library-multifunctional capture probe module complex to remove the single stranded 3′ ends.

In some examples, the enzyme for use in the 3′-5′ exonuclease enzymatic processing is T4 DNA polymerase.

In certain examples, the methods contemplated herein comprise performing PCR on the 3′-5′ exonuclease enzymatically processed complex, wherein a tail portion of the multifunctional capture probe molecule is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the cDNA target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence.

In further examples, a method for targeted gene expression analysis is provided comprising: (a) hybridizing a tagged RNA expression library with a multifunctional capture probe module complex, wherein the multifunctional capture probe module selectively hybridizes to a specific target region in the tagged RNA expression library; (b) isolating the tagged RNA expression library-multifunctional capture probe module complex from a); (c) performing 3′-5′ exonuclease enzymatic processing on the isolated tagged RNA expression library-multifunctional capture probe module complex from b) using an enzyme with 3′-5′ exonuclease activity to remove the single stranded 3′ ends; (d) performing PCR on the enzymatically processed complex

from c) wherein the tail portion of the multifunctional capture probe molecule is copied in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the target region capable of hybridizing to the multifunctional capture probe module and the complement of the multifunctional capture probe module tail sequence; and (e) performing targeted gene expression analysis on the hybrid nucleic acid molecule from d).

In additional examples, a method for targeted gene expression analysis is provided comprising: (a) hybridizing a tagged RNA expression library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the RNA expression library; (b) isolating the tagged RNA expression library-multifunctional capture probe hybrid module complex from a); (c) performing 5' to 3' DNA polymerase extension of the multifunctional capture probe on the complex from b) to replicate a region of the captured, tagged target region that is 3' of the multifunctional capture probe in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of the tagged target region that is located in the 3'direction from the location where the multifunctional capture probe hybrid module hybridizes to the target region; and (d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In various examples, a method for targeted gene expression analysis is provided comprising: (a) hybridizing a tagged cDNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the cDNA library; (b) isolating the tagged cDNA library-multifunctional capture probe hybrid module complex from a); (c) performing 5' to 3' DNA polymerase extension of the multifunctional capture probe on the complex from b) to replicate a region of the captured, tagged target region in the cDNA library that is 3' of the multifunctional capture probe in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the multifunctional capture probe hybrid module and the complement of the tagged target region in the cDNA library that is located in the

3'direction from the location where the multifunctional capture probe hybrid module hybridizes to the target region; and (d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In particular examples, at least two different multifunctional capture probe modules are used in the at least two (a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In certain examples, at least one multifunctional capture probe module hybridizes downstream of the target region and at least one multifunctional capture probe module hybridizes upstream of the target region.

In additional examples, a method for cDNA sequence analysis is provided comprising: (a) obtaining one or more clones, each clone comprising a first cDNA sequence and a second cDNA sequence, wherein the first cDNA sequence comprises targeted genomic cDNA sequence and the second cDNA sequence comprises a capture probe sequence; (b) performing a paired end sequencing reaction on the one or more clones and obtaining one or more sequencing reads; and (c) ordering or clustering the sequencing reads of the one or more clones according to the probe sequence of the sequencing reads.

In various examples, a method for cDNA sequence analysis is provided comprising: (a) obtaining one or more clones, each clone comprising a first cDNA sequence and a second cDNA sequence, wherein the first cDNA sequence comprises targeted genomic DNA sequence and the second cDNA sequence comprises a capture probe sequence; (b) performing a sequencing reaction on the one or more clones in which a single long sequencing read of greater than about 100 nucleotides is obtained, wherein the read is sufficient to identify both the first cDNA sequence and the second cDNA sequence; and (c) ordering or clustering sequencing reads of the one or more clones according to the probe sequence of the sequencing reads.

In particular examples, the methods contemplated herein comprise analyzing all sequencing reads associated with a second read sequence by: determining the distributions of unique and redundant sequencing reads; counting the number of times a unique read is encountered; fitting a frequency distribution of the unique reads to a statistical distribution; inferring a total number of unique reads; and converting

unique read counts into transcript abundance using normalization to the total reads collected within each cDNA library sample.

In certain examples, a method for targeted genetic analysis is provided comprising: (a) hybridizing a tagged DNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the DNA library; (b) isolating the tagged DNA library-multifunctional capture probe hybrid module complex from a); (c) performing a concerted enzymatic processing of the tagged DNA library-multifunctional capture probe hybrid module complex from b) that includes 5′ FLAP endonuclease activity, 5′ to 3′ DNA polymerase extension, and nick closure by a DNA ligase to join the complement of the multifunctional capture probe to the target region that is 5′ of the multifunctional capture probe binding site in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the complement of the multifunctional capture probe hybrid module and a region of the tagged target region that is located 5′ of the location where the multifunctional capture probe hybrid module hybridizes to the genomic target region; and (d) performing targeted genetic analysis on the hybrid nucleic acid molecule from c).

In various examples, steps a) through c) are repeated at least about twice and the targeted genetic analysis of d) comprises a sequence alignment of the hybrid nucleic acid molecule sequences obtained from the at least two d) steps.

In certain examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In particular examples, at least one multifunctional capture probe module hybridizes downstream of the target region and at least one multifunctional capture probe module hybridizes upstream of the target region.

In additional examples, a method for determining copy number of a specific target region is provided comprising: (a) hybridizing a tagged DNA library with a multifunctional capture probe hybrid module complex, wherein the multifunctional capture probe hybrid module selectively hybridizes to a specific target region in the genomic library; (b) isolating the tagged DNA library-multifunctional capture probe

hybrid module complex from a); (c) performing a concerted enzymatic processing of the tagged DNA library-multifunctional capture probe hybrid module complex from b) that includes 5′ FLAP endonuclease activity, 5′ to 3′ DNA polymerase extension, and nick closure by a DNA ligase to join the complement of the multifunctional capture probe to the target region that is 5′ of the multifunctional capture probe binding site in order to generate a hybrid nucleic acid molecule, wherein the hybrid nucleic acid molecule comprises the complement of the multifunctional capture probe hybrid module and a region of the tagged target region that is located 5′ of the location where the multifunctional capture probe hybrid module hybridizes to the target region; and (d) performing PCR amplification of the hybrid nucleic acid molecule in c); and (e) quantitating the PCR reaction in d), wherein the quantitation allows for a determination of copy number of the specific target region.

In various examples, the methods contemplated herein comprise obtaining the sequences of the hybrid nucleic acid molecules from step d).

In particular examples, steps a) through d) are repeated at least about twice and a sequence alignment of the hybrid nucleic acid molecules from the at least two d) steps.

In particular examples, at least two different multifunctional capture probe modules are used in the at least two a) steps, wherein the at least two a) steps employ one multifunctional capture probe module each.

In certain examples, at least one multifunctional capture probe module hybridizes downstream of the genomic target region and at least one multifunctional capture probe module hybridizes upstream of the genomic target region.

In additional examples, the targeted genetic analysis is sequence analysis.

In further examples, the target region is a genomic target region and the DNA library is a genomic DNA library.

In some examples, the target region is a cDNA target region and the DNA library is a cDNA library.

In one aspect, the present disclosure refers to a complex comprising: (a) a tagged DNA molecule comprising a DNA fragment, wherein each end of the DNA fragment is ligated to a multifunctional adaptor module comprising: (i) a first region comprising a random nucleotide tag sequence; (ii) a second region comprising a sample code sequence; and (iii) a third region comprising a PCR primer sequence; and (b) a hybrid nucleic acid molecule hybridized to the tagged DNA molecule, wherein the hybrid nucleic acid molecule comprises, from 5' to 3': (i) a tail sequence comprising a PCR primer binding site; (ii) a capture probe to which a portion of the DNA fragment is hybridized; and (iii) the complement of a region of the DNA molecule that is located 3' to the capture probe, wherein the region comprises a multifunctional adaptor module.

In one aspect, the present disclosure refers to a complex comprising: (a) a multifunctional capture probe module comprising, from 3' to 5': (i) a first region capable of hybridizing to a partner oligonucleotide; (ii) a second region capable of hybridizing to a specific genomic target region in a DNA fragment; and (iii) a third region comprising a tail sequence; and (b) a tagged DNA molecule comprising the DNA fragment, wherein the tagged DNA molecule is hybridized to the multifunctional capture probe module, wherein the 5' end of the DNA fragment is ligated to a multifunctional adaptor module comprising: (i) a first region comprising a random nucleotide tag sequence; (ii) a second region comprising a sample code sequence; and (iii) a third region comprising a PCR primer sequence; wherein the 3' end of the DNA molecule comprises the compliment of the third region of the multifunctional capture probe module.

In one aspect, the present disclosure refers to a complex comprising: (a) a tagged DNA molecule comprising a DNA fragment, wherein each end of the DNA fragment is ligated to a multifunctional adaptor module; wherein the multifunctional adaptor module comprises: (i) a first region comprising a random nucleotide tag sequence; (ii) a second region comprising a sample code sequence; and (iii) a third region comprising a PCR primer sequence; and (b) a multifunctional capture probe module comprising: (i) a first region capable of hybridizing to a partner oligonucleotide; and (ii) a second region capable of hybridizing to a specific target region in the DNA fragment; wherein the tagged DNA molecule is hybridized to the multifunctional capture probe module.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Figure 1: Construction of an amplifiable, sample coded, tagged genomic DNA library. Purified genomic DNA was isolated from a source such as whole blood or a buccal cheek swab. The DNA was fragmented (e.g. by mechanical, enzymatic or chemical means) and the ends of the DNA were repaired, in this example, to blunt ends. The repaired DNA was ligated to a multifunctional adaptor module that contained a universal amplification sequence, a random nucleotide tag sequence and a sample code sequence. A specific example of a typical adaptor duplex molecule is shown by way of example.

Figure 2: Genomic capture probe design. (A) Structure of a prototypical 114 nt probe. Region 1 is a 34 nt sequence that shares perfect complementarity to a highly modified, 34 nt partner oligonucleotide. The partner oligonucleotide was modified on its 5′ end with a biotin – TEG chemical entity that enabled capture on streptavidin-coated magnetic beads. B stands for the "Bio-TEG" modification. Region 2 is the 60 nt probe region designed to interact with genomic DNA targets. Region 3 is a 20 nt tail that introduced PCR amplification sequences into the captured genomic fragments. (B) An example of the highly modified partner strand that is complementary to region 1 of each individual probe.

Figure 3: Sequence "spreading" in conventional, hybridization-based capture experiments. (A) Flanking fragments can "hitch-hike" into capture libraries by way of legitimate hybridization triplexes involving probe-fragment-flanking fragment interactions. (B) The net result of cross-fragment hybridization is sequence "spreading", meaning sequences that are within proximity to the target region (dashed lines) but fall outside the boundaries of the desired target.

Figure 4: Enzymatic processing of fragment: probe hybridized complexes. (A) The purified complex of fragment (light gray) and probe (black) prior to processing. B – biotin affinity modification. (B) DNA polymerase (e.g. T4 DNA polymerase) encoded 3′ → 5′ exonuclease activity removes the 3′ segment of the captured fragment. (C) Upon encountering the probe: fragment duplex region, the polymerase copies the probe tail segment onto the hybridized genomic fragment. (D) The final modified